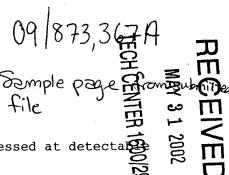
Normal Colon



Genes that are expressed in normal colon, that are not expressed at detectal levels in colon adenocarcinoma

Below is a listing of those genes that are expressed at appreciable levels in normal colon, but that do not appear to be expressed in colon adenocarcinoma. There are 333 sequences presented in the listing below.

>gi|1472311|gb|AA011199.1|AA011199 ze23c03.s1 Soares\_fetal\_heart\_NbHH19W Homo sapiens cDNA clone

IMAGE: 359812 3', mRNA sequence

>gi|1479353|gb|AA016979.1|AA016979 ze41h01.s1 Soares retina N2b4HR Homo sapiens cDNA clone IMAGE:361585 3' similar to gb|M10329|MUSUR48S Mouse 4.8S U6 small

nuclear (rRNA); contains Alu repetitive element;, mRNA sequence AAATATGGAACGCTTCACGAATTTGCGTGTCATCCTTGCGCAGGGGCCATGCTAATCTTCTCTGTATCGT TCCAATTTTAGTATATGTGCTGCCGAAGCGAGCACCGTGCTTAGTTATTCTAAGTGAGGGCCCCAGGATC CACCTGCCTAGGCTTCCCAAAGTGCTGGGATTACAGGCGTGACCCACCGCGCCCAGCCAAGTTTTGGTTT CCTCAACTGGAGGTAATATTACATATTTTACTTATACATATGCATAAGTAAACAAAGAGGGTTGTTTTGA GGGTCAAATAAATTGATGGATGTTAACGCTCTNCTGGTAAATTATAAAGCACTATACAAATACAAGGCAT TATTGTTAATAATAGAGCTTAATTACACCTGTCCTCATTTGATCTCTCANAGACC

>gi|1493220|gb|AA027011.1|AA027011 zk02c08.sl Soares\_pregnant\_uterus\_NbHPU Homo sapiens cDNA clone

IMAGE: 469358 3', mRNA sequence

>gi|1512487|gb|AA037388.1|AA037388 zc03e01.s1

Soares\_parathyroid\_tumor\_NbHPA Homo sapiens cDNA clone IMAGE:321240 3', mRNA sequence

## Sample Sequence listing

```
Smith, John: Smithgene Inc.
 <110>
              Example of a Sequence Listing
 <120>
              01-00001
 <130>
                                                                                ٠
              PCT/EP98/00001
 <1(0)
 <1(1)
              1998-12-31
              US 08/999,999
 <150>
              1997-10-15
 <151>
 <160>
              PatentIn version 2.0
 <170>
 <210>
              389
<211>
 <212>
              NAG
              Paramecium sp.
<213>
<220>
              CDS
<221>
              (279) . . . (389)
<222>
<300>
<301>
             Doc. Richard
              Isolation and Characterization of a Gene Encoding a
<302>
             Protease from Paramecium sp.
<303>
             Journal of Genes
<304>
<305>
             4
<306>
             1 - 7
<307>
             1988-06-31
<308>
             123456
<309>
             1988-06-31
<400>
                                                                                        60
                                       cigggetici
                                                    caccetgeta
                                                                 atcagatete
             attectgtgt
                         cctcttctct
agciglagic
                                                                                E
                                                                                       120
                                                    caggeaggea
                                                                 ggcaggcagc
                                       tgcagcttca
agggagagtg
             tottgaccèt:
                          cctctgcctt
                                                                                       180
                                                    aggettaggg
                                                                 tgggttccgc
                                       ctttcagcc
rgargrggca
             attigctggca
                          gtgccacagg
                                                                                      240
                                                                 cctctcgctc
                                                    ctctcgctct
                          cgcgctcctc
                                       tcgcgcctct
cgcggcgcgg
             cggcccctct
```

ゴ

::**::**:::

## Appendix 3, page 2

```
ttc
                                                                         atg
                                                              gtt
                                                                   tca
                                                         atg
                                                                                          296
                                          cagttage
                            9499499999
               aggtgagcag
                                                              Val Ser Het
                                                                              Phe
                                                                                    Ser
                                                         Het
                                                           1
                                                         ttt gtt
                                                                              ttc
                                              tgt ttg
                                                                   tgt
                                                                         ttg
                       tgg cct gga
Trp Pro Gly
                                        ttt
                                                                                    CAA
                                                                                          344
             ttc
                  888
 ttg
       tct
                                                                              Phe
                                                        Phe
                                                              Val
                                                                    Cys
                                                                         Leu
                                                                                    Cln
                                              Cys. Lcu
                                        Phe
                  Lys Trp
       Ser
             Phe
 Lcu
                                                                          20
                                               15
                   10
                                                                  cco aat ctt
                                                        ctg cag
                                                                                          389
                      ctc ccc tgt
Leu Pro Cys
                                        CAC
                                              tca
                                                  tca
                 gtc
 tgt
       CCC
             888
                                                                        Asn. ...biu
                                                       -Leu Gln
                                                                   Pro
                                        llis
                                              Ser Ser
             Lys -Val
 Cys
       Pro
                                                                    35
                                         30
              25
                                                                            يو التير
                                                                                Jen.
               2
<210>
               37
 <211>
 <212>
               LBJ.
 <213>
               Paramecium sp.
 <000>
                                                             Pro
                                                                   Gly Phg
                                                                            Cys
                                                       Trp
                                            Phe
                                                 Lys
                                       Scr
            Ser Het Phe
                            Ser
                                  Lcu
      Val
 Het
                                                                                   Ser -
                                                                   Cys
                                                             rro
                                                                              Scr
                                                       Leu
                                       Pro
                                                  Val
                       Phe
                            Cln
                                  Cys
      Val
                  Lcu
                                                                         30
                  20
            Pro
 Leu
      Gln
                 ۸sn
                       LQU
             3.50
              ì
<210>
<211>
              11
              PRT
 <212>
              Artificial Sequence
 <213>
<220>
              Designed pentide based on size and polarity to act as a
<223>
              linker between the alpha and beta chains of Protein XYZ.
<400>
                           Pro Mct His Thr Glu
                                                       lle
                      Glu
Het Val
           Λsn
                 Leu
                                                  10
<210>
<400>
000
```

[Annex VIII follows]

E

....

table. The numeric ide fier shall be used only in the equence listing." The order and presentation of the items of in mation in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a n w line and shall begin with the item of information shall begin on a n w line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequenc Listing." The following table illustrates the numeric identifiers.

٠.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	Preferably max. of 10 names; one name per line; preferable format: Surname, Other. Names and/or Initials	M v
<120>	Title of Invention	~ <del>~</del>	<b>м</b> Детек
<130>	File Reference	Personal file reference	M, when filed prior to assignment of appl. number
<140>	Current Applica- tion Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	И
<170>	Software	Name of software used to create the Sequence Listing	0
<210>	SEQ ID NO: #:	Response shall be an integer representing the SEQ ID NO shown	М
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M

... Whether presented sequence moleculc is DNA, RNA, or PRT (protein). If . a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/ NNV wolccale shall be further described in the <220> to <223> [cature section.

<213>

Organism.

Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.

<220>

Feature

Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.

M, under the following conditions: if "n,"
"Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGAN-ISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.

<221>

Name/Key

Provide appropriate identifier for feature, pre(crably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6

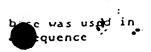
M, under the following conditions:= if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence

<222>

Location

Specify location within sequence; where appropriate state number of first and last bases/amino acids

M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified



<223> Other Information

Other relevant information; four lines maximum

<300> ;	Publication Information	Leave blank O after <300> /	
<301>	Autho <i>rs</i>	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	
<302>	Title	o	
<303>	Journal	0	
<304>	Volume	, <b>0</b>	
<305>	Issue	, o	
<306>	Pages	0 ·	
<307>	Date	Journal date on which O data published; specify as yyyy-mm-dd, 1001-yyyy or Season-yyyy	
<300>	Database Accession Number	Accession number O assigned by data- base including database name	
<309>	Database Entry Date	Date of entry in Odatabase; specify as yyyy-mm-dd or MMM-yyyy	
<310>	Patent Document Number	Document number; O for patent-type citations only. Specify as, for example, US 07/999,999	

t-

Patent File Document filing
date, for patenttype citations only;
specify as yyyy-mm-dd

date, for patent-type citations only; specify as yyyy-mm-dd

O

<313> Relevant FROM (position) TO 0 Residues (position)

<400> Sequence SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual

5. Section 1.024 is revised to read as follows:

- 1.024 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.
- (a) The computer readable (orm required by 1.021(c) shall meet the following specifications:

sequence

- (1) The computer readable form shall contain a single: "Sequence Listing" as either a diskette, series of diskettes, or other permissible media: outlined in paragraph (c) of this section.
- (2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.
- (3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.
- (4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.
- (5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" (ile.
- (6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.
- (b) Computer readable form submissions must meet these format requirements:
- (1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;
- (2) Operating System: MS-DOS, Unix or Macintosh: